

INFO323 – THÈME 3

DONNÉES MANQUANTES

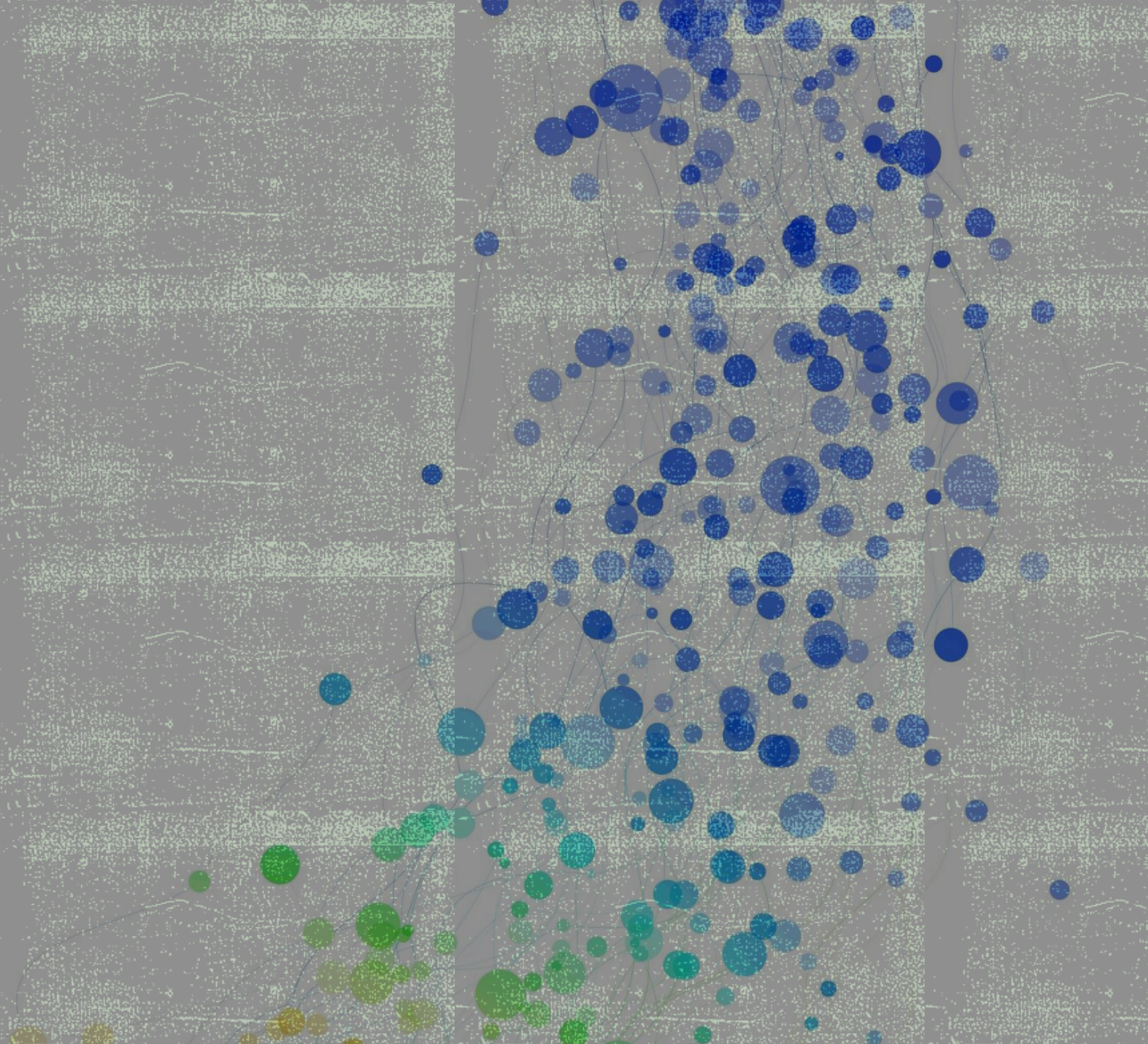
Nadia Tahiri, Ph. D.
Professeure adjointe
Université de Sherbrooke

Nadia.Tahiri@USherbrooke.ca



Source: Luc Lavoie (UdeS) et Christina Khnaisser , BD013

ORIGINE





L'information est applicable, mais la valeur n'est pas encore connue (*date de décès d'une personne vivante*).



L'information est inapplicable (*nombre de sommets d'un cercle*).



L'information existe, mais il n'est pas permis (*légalement*) de l'enregistrer (*religion d'un employé*).



L'information existe, mais on n'a pas les moyens de trouver la valeur (*évaluation d'un employé alors qu'il travaillait pour une organisation concurrente*).



L'information existe, mais elle n'est pas encore enregistrée (*en raison de l'absence de l'employé préposé à la saisie*).

CAUSALITÉ



L'information est enregistrée, mais pas encore disponible (*texte écrit, saisi, stocké, mais pas encore publié*).



L'information a été enregistrée puis supprimée (*un utilisateur ne veut plus que le nom de son conjoint soit conservé*)



L'information est disponible, mais en changement et donc potentiellement invalide (*solde d'un compte bancaire sur lequel une opération est en cours*).



L'information est disponible, mais on ne sait pas si elle est fiable (*la note d'examen non encore approuvée par le doyen*).



L'information est disponible, mais invalide (*si une erreur s'est produite lors du calcul de la valeur*)

CAUSALITÉ



CAUSALITÉ

- La classe d'information est sécurisée (*les informations personnelles des professeurs ne sont pas accessibles aux étudiants*).
- L'objet représentant l'information est sécurisé (*un utilisateur bloque l'accès à ses infos personnelles sur un réseau social*).
- Une information est sécurisée durant un certain laps de temps (*le budget préalablement à sa communication au parlement*).
- L'information est calculée à partir d'au moins une information absente ou incertaine (*l'âge en fonction d'une date de naissance par ailleurs absente*).



GESTION



Incomplet.

En fait, les raisons de l'absence varie selon le contexte et leur interprétation selon la finalité de la requête.



Trop complexe.

Tant pour la saisie que pour les requêtes, la complexité induite est trop élevée en regard de de l'effort requis et des moyens de vérification/validation.

UN MODÈLE SIMPLE (PROPOSÉ PAR CODD)

- **N**
 - L'information n'est **pas applicable**.
 - Dans ce cas, l'utilisation de l'annulabilité est à remettre en question; une bonne modélisation permet généralement d'éviter d'y avoir recours.
- **I**
 - L'information est **inconnue**.
 - Dans ce cas, l'annulabilité pourrait être légitime; la question est de savoir comment la représenter pour que cela pose le moins de problèmes possible.
- **X**
 - L'information n'est **pas accessible**.
 - **À court terme** : le gestionnaire transactionnel permet d'éviter l'utilisation de l'annulabilité en différant la mise à disponibilité tout en conservant le contrôle des accès concurrents et en préservant la cohérence de la BD.
 - **À long terme** : équivalent à **I**.





SOLUTIONS AVEC ANNULABILITÉ

- Que faire lorsqu'une donnée est absente?
- Trois solutions classiques
 - corriger cette lacune à la source (dans la réalité);
 - modifier le modèle pour en tenir compte;
 - introduire la notion d'*annulabilité* dans la théorie relationnelle, ce qui induit le recours
 - à une logique non classique
 - (afin de pouvoir définir l'égalité, essentielle aux opérations d'affectation, de restriction, de jointure, d'union...)
 - et à l'un des deux artifices suivants :
 - un marqueur NUL (une propriété des attributs) ou
 - une valeur NULLE (ajoutée à tous les domaines).





SOLUTIONS SANS ANNULABILITÉ

- Principes
 - Séparer les propositions complètes des incomplètes.
 - Conserver les causes d'absence séparément.



4V - BELNAP

f_{\neg}		f_{\wedge}	T	B	N	F	f_{\vee}	T	B	N	F
T	F	T	T	B	N	F	T	T	T	T	T
B	B	B	B	B	F	F	B	T	B	T	B
N	N	N	N	F	N	F	N	T	T	N	N
F	T	F	F	F	F	F	F	T	B	N	F

B : sur-déterminé;
N : sous-déterminé

C'était la proposition révisée de Codd... et elle n'a pas été suivie par le comité de standardisation du langage SQL.

LOGIQUE NON CLASSIQUE



3V-KLEENE

P3 (de Priest : – I est sur-déterminée – T et I sont vraies – avec tautologie)

\neg		\wedge	T	I	F	\vee	T	I	F	\rightarrow	T	I	F	\leftrightarrow	T	I	F
T	F	T	T	I	F	T	T	T	T	T	T	I	F	T	T	I	F
I	I	I	I	I	F	I	T	I	I	I	T	I	I	I	I	I	I
F	T	F	F	F	F	F	T	I	F	F	T	T	T	F	F	I	T

B3 (faible : I est réductrice – T seule valeur vraie – avec tautologie)

\neg		\wedge	T	I	F	\vee	T	I	F	\rightarrow	T	I	F	\leftrightarrow	T	I	F
T	F	T	T	I	F	T	T	I	T	T	T	I	F	T	T	I	F
I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I
F	T	F	F	I	F	F	T	I	F	F	T	I	T	F	F	I	T

K3 (forte : I est sous-déterminée - T seule valeur vraie – pas de tautologie)

non retenue

LOGIQUE NON CLASSIQUE

